Analysis of Skewed Data by using Compound Poisson-Exponential Distribution with Applications to Insurance Claims

Abstract

The main aim of this paper is to introduce a new family of distributions, namely compound zero-truncated Poisson exponential distribution of which exponential distribution is a special case. The proposed family of distributions represents the zero truncated-Poisson sum of independent and identically distributed exponential random variables. The proposed distribution has two parameters and its probability density function can be skewed and unimodal. It can be used quite effectively in analyzing skewed data. The maximum likelihood estimators of the unknown parameters cannot be obtained in closed forms. For this reason, we suggest to use EM type algorithm to estimate the unknown parameters, and it is observed that it is easy to implement in practice. We further consider the bivariate version of the proposed model which has three parameters and provides different properties. The EM type algorithm, likewise, can be used to estimate unknown parameters. We have performed an extensive simulation studies to see the performances of the proposed EM algorithm, and a real data set has been analyzed to see the effectiveness of the proposed models. The performances are quite satisfactory.

Keywords: Bivariate distribution; EM type algorithm; exponential variables; Fisher information matrix; maximum likelihood estimators; zero truncated Poisson distribution.

AMS 2012 Subject Classifications: 62E15; 62F10; 62H10.

1 Introduction

Componding technique has been used quite effectively during the last few decades, in developing new probability distributions. The compound distributions have many natural applications in actuarial science, insurance, risk model, financial economics and finance. For example, a compound distribution may be used to model the aggregate claims during a fixed policy period for an insurance policy, or to model the aggregate claims during a fixed policy period for a group of independent insured individuals. In other words, it is desirable to discuss distributions that can either model the total claims for an insured individual or a group of independent insured individuals over a fixed period such that the claim frequency is uncertain (no claim, one claim or multiple claims).

Two compound distributions are of special importance in finance and insurance. The compound Poisson distribution is often a popular choice for aggregate claims modeling because of its desirable properties. The computational advantage of the compound Poisson distribution enables us to easily evaluate the aggregate claims distribution when there is at least one claim. It is well documented in the literature that the exponential distribution plays a pivotal role for modeling the claim severity in actuarial science and it can be used quite effectively to analyze skewed data. Following the work of Azzalini [5] on skew normal distribution, many researchers introduced different skewed distributions, see for example Asgharzadeh et al. [4], Wahed and Ali [28]), Ma and Genton [19], Barreto-Souza [7], Chakraborty et al. [11], Asgharzadeh et al. ([2], [3]) and see the references cited therein. In other contexts, several compound distributions have been introduced in the literature by different authors, see for example Bakouch et al. [6], Rodrigues, et al. [26], Mahmoudi and Jafari [20], Kundu [16], Kozubowski and Panorska [15], Revfeim [25] etc.

Compound models have been used quite extensively in analyzing insurance losses data. In this context, extensions of the two-parameter Birnbaum-Saunders (BS) model (Birnbaum and Saunders [10]) have been developed. One approach for extending the BS model is to replace the standard normal variable in the stochastic representation of the BS model with other variables followed by skewed distributions. Hashemi et al. [14] proposed the normal mean-variance Lindley BS (NMVL-BS) distribution and used it for modeling the financial data. Another generalization is the normal mean-variance generalized BS (NMV-GBS) distribution (cf. Naderi et al. [23]) which can effectively analyze economic real data. A general approach to compound unimodal hump-shaped distributions with a mixing dichotomous distribution is discussed in Tomarchio and Punzo [27]. They used the models in fitting real insurance loss data. Bernardi et al. [8] have considered the skew exponential power (SEP) distribution for modeling the Bayesian quantile regression. It is shown that their framework produces more reliable estimates for heavy-tailed data. Composite models and their applications in the insurance losses can also be found in Grun and Miljkovic [13] and Bhati and Ravi [9].

In studying the insurance claims, it is extremely important to consider the number of claims, N, as well as the corresponding payments in marginal and bivariate forms. One important point to note that not all claims lead to payment by an insurance company, i.e. some of the claims are not approved. Hence, the sum of approved payment amounts are recorded. Based on this concept, Raqab et al. [24] proposed a new distribution as compound zero-truncated Poisson skew normal distribution. This model is based on the random sum of N independent Gaussian random variables. In this work, we develop another version of skewed distribution based on modeling aggregate claims. Let N and L^* (conditioning that $L^* \geq 1$) denote the total number of claims and approved claims, respectively, in a fixed time period, and $T_1, \ldots, T_{L^*}$ denote the amount paid by the insurance company. Hence, $X = T_1 + \ldots + T_{L^*}$, denotes the aggregate amount paid by the insurance company in that fixed time period. It is assumed that T_i's are independent identically distributed (i.i.d.) exponential random variables, and L^* is a zero truncated Poisson random variable and it is independent of T_i's. We call this new distribution as a *compound zero-truncated Poisson exponential (ZTP-EXP) distribution.* Moreover, the exponential distribution can be obtained as a special case of the proposed compound ZTP-EXP distribution.

Additionally, we extend the proposed model to bivariate case with three parameters and we refer to this new distribution as the *bivariate zero-truncated Poisson exponential model (BZTP-EXP).* The BZTP-EXP model can be used as appropriate model for many applications in climatology, actuarial sciences, water and finance. We here concentrate in finance application and use the proposed model for analyzing the total of insurance claimed corresponding to the number of claimants for a given period of time. The ZTP-EXP distribution has two parameters, whereas the BZTP-EXP distribution has three parameters.

Estimation of the unknown parameters is always a challenge problem in any data analysis. In this case it is observed that the maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in closed forms. They can be acquired by solving higher dimensional optimization problems. To avoid that we have suggested an effective *expectation-maximization* (EM) type algorithm to compute the MLEs of unknown parameters, see Kundu and Nekoukhou [17] in case of discrete missing values. It is observed that at each 'E'-step, the corresponding 'M'-step can be performed explicitly. We further provide an expression for the observed Fisher information matrix as suggested by Louis [18] to construct the confidence intervals for unknown parameters. We present some simulation study to show the effectiveness of the proposed EM type algorithm. For illustrative purposes, we analyze two real data sets using the proposed models.

The rest of the paper is organized as follows. The formulation of the proposed univariate compound distribution (ZTP-EXP) and discussion of some general properties are presented in Section 2. In Section 3, the BZTP-EXP model is introduced and all the necessary theoretical results and general properties are provided. Also, a study of estimation and inference statistical of BZTP-EXP distribution parameters are introduced in this section. Some simulation results and the analysis of two real data sets are presented in Section 4. We provide some goodness-of-fit measures and tail-risk measures and show how it can be used to choose the best fitted model, in Section 5 and finally we conclude the paper in Section 6.

2 Univariate Compound ZTP-EXP Distribution

In this section, we introduce the proposed compound ZTP-EXP and provide all the necessary theoretical results and general properties.

2.1 Description and Properties

Suppose $\{T_1, T_2, ...\}$ is a sequence of i.i.d. exponential random variables with probability density function (PDF)

$$f(t, \beta) = \beta e^{-\beta t} \; t > 0, \; \beta > 0, \tag{1}$$

the random variable N has Poisson distribution with parameter λ and it is independent of $\{T_i : \; i = 1, 2, ...\}$. Define a random variable L as follows

$$L = \sum_{i=1}^{N} I_i = I_1 + I_2 + ... + I_N, \tag{2}$$

where $I_i = 0$ implies that the insurance company has not paid the claim, whereas $I_i = 1$ implies that the insurance company has paid the claim. Note that, $P(I_i = 1) = 1 - P(I_i = 0) = p$, for $0 \leq p \leq 1$. Here L denotes the total number of claims which have been paid by the insurance company. The distribution of $\{L|N = n\}$ is a binomial with parameters n and p. Hence, L is Poisson distribution with mean $\theta = \lambda p$. Now, let us consider

$$L^{\star} \stackrel{D}{=} \{L \mid L \geq 1\}, \tag{3}$$

where $\stackrel{D}{=}$ denotes equal in distribution, and it independent of $\{T_1, T_2, ...\}$. Then, $L^\star$ has the following probability mass function (PMF)

$$P(L^\star = l) = \frac{e^{-\theta}\theta^l}{l!(1-e^{-\theta})}, \; l = 1, 2, 3, ..., \quad \theta > 0. \tag{4}$$

Thus, the random variable $L^\star$ has a zero-truncated Poisson distribution with parameter θ. The random variable X is said to follow a ZTP-EXP distribution if it can be represented by

$$X = \sum_{i=1}^{L^\star} T_i. \tag{5}$$

From now on, we write $X \sim$ ZTP-EXP(β, θ). Note that

$$P(X \leq x, L^\star = l) = \frac{e^{-\theta}\theta^l}{l!(1-e^{-\theta})} \cdot \frac{\Gamma_{\beta x}(l)}{\Gamma(l)},$$

where $\Gamma_x(a) = \int_0^x t^{a-1}\, e^{-t} dt$, is the incomplete gamma function with $a > 0$. Hence, the joint PDF of X and $L^\star$ is given by

$$f_{X,L^\star}(x, l) = \frac{e^{-\theta}\theta^l}{l!(1-e^{-\theta})} \cdot \frac{\beta e^{-\beta x}(\beta x)^{l-1}}{\Gamma(l)}, \; x > 0, \quad l = 1, 2, ..., \tag{6}$$

It can be easily checked that the exponential distribution with parameter β is a special case when $\theta \to 0$ and $l = 1$. Specifically,

$$\lim_{\theta \to 0} P(X \leq x, L^\star = l) = \begin{cases} 1 - e^{-\beta x} & \text{if } l = 1, \\ 0 & \text{if } l > 1. \end{cases}$$

The cumulative distribution function (CDF) of X can be obtained as follows:

$$\begin{aligned} P(X \leq x) &= \sum_{l=1}^{\infty} P(X \leq x, L^* = l) \\ &= \frac{e^{-\theta}}{1-e^{-\theta}} \sum_{l=1}^{\infty} \frac{\Gamma_{\beta x}(l)}{\Gamma(l)} \cdot \frac{\theta^l}{l!}. \end{aligned} \tag{7}$$

The corresponding PDF of X can be written as

$$f_X(x) = \frac{\beta e^{-\beta x} e^{-\theta}}{1-e^{-\theta}} \sum_{l=1}^{\infty} \frac{(\beta x)^{l-1}\theta^l}{l!\; \Gamma(l)} \;\; , \; x > 0; \; \beta, \theta > 0. \tag{8}$$

From Eq.(8), the parameter β controls the scale of the distribution while the parameter θ controls its shape. Clearly, the variate X is an exponential distribution with mean $1/\beta$ when θ tends to be 0. The PDFs of ZTP-EXP for different values of the parameters have been displayed in Figure 2.1. It can take variety of shapes. The PDF of X is always unimodal and positively skewed. Also it is noticed that as θ increases, the skewness increases.

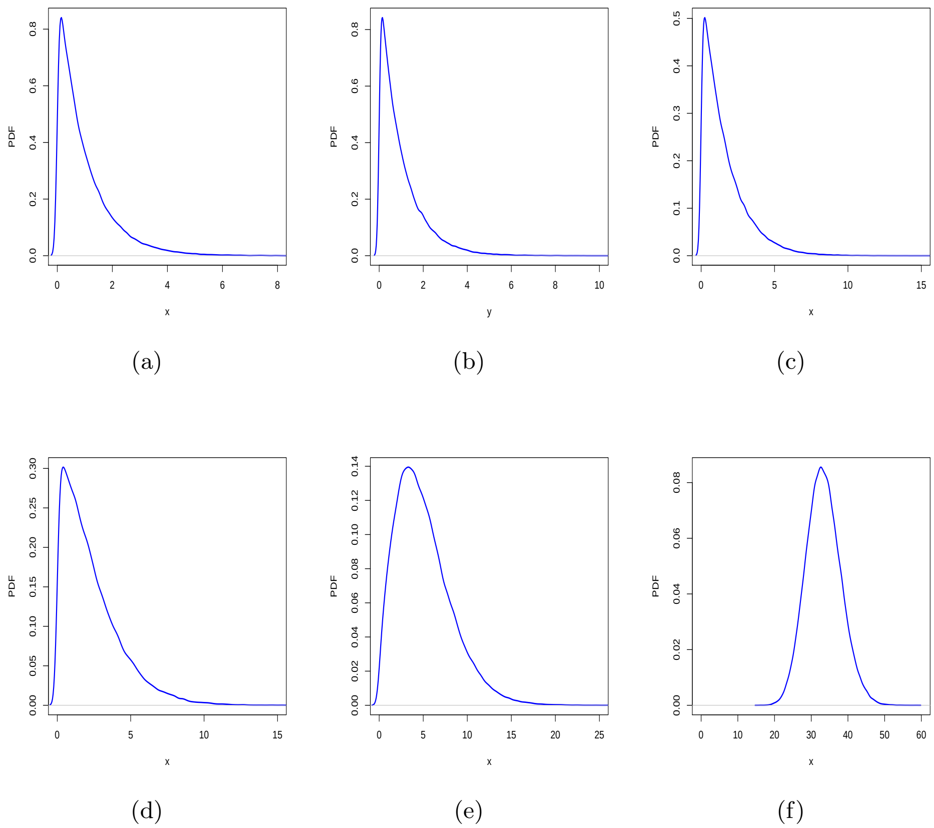

Figure 1: PDF plots of ZTP-EXP (β, θ) distribution for different parameter values: (a) (1, 0.0001) (b) (1, 0.005) (c) (1, 1) (d) (1, 2) (e) (1, 5) (f) (3, 100).

Now, consider $(X, L^\star)$ which has the joint PDF as given in (6). The conditional CDF of $(X, L^\star)$ given $L^\star$ is

$$P(X \leq x, L^\star \leq n \mid L^\star \leq l) = \begin{cases} \dfrac{1}{\sum_{j=1}^{l} \theta^j/j!} \displaystyle\sum_{k=1}^{n} \dfrac{\Gamma_{\beta x}(k)\ \theta^k}{\Gamma(k)\ k!} & \text{if } l \geq n > 0, \\[2ex] \dfrac{1}{\sum_{j=1}^{l} \theta^j/j!} \displaystyle\sum_{k=1}^{l} \dfrac{\Gamma_{\beta x}(k)\ \theta^k}{\Gamma(k)\ k!} & \text{if } l < n. \end{cases}$$

Therefore, for any integer $n > 0$ and $y \geq x$, the conditional CDF of $(X, L^\star)$ given $X \leq y$ is

$$\begin{aligned} P(X \leq x, L^\star \leq n \mid X \leq y) &= \frac{P(X \leq x, L^\star \leq n)}{P(X \leq y)} \\ &= \frac{1}{\sum_{j=1}^{\infty} \frac{\Gamma_{\beta x}(j)\ \theta^j}{\Gamma(j)\ j!}} \sum_{k=1}^{n} \frac{\Gamma_{\beta x}(k)\ \theta^k}{\Gamma(k)\ k!}. \end{aligned} \tag{9}$$

From (9), we further have

$$P(L^\star \leq l \mid X \leq x) = \frac{1}{\sum_{j=1}^{\infty} \dfrac{\Gamma_{\beta x}(j)\ \theta^j}{\Gamma(j)\ j!}} \sum_{k=1}^{l} \frac{\Gamma_{\beta x}(k)\ \theta^k}{\Gamma(k)\ k!}. \tag{10}$$

Now, the conditional PMF of $L^\star$ given $X = x$ is given by

$$P(L^\star = l \mid X = x) = \frac{\dfrac{(\theta\ \beta\ y)^l}{l!\ \Gamma(l)}}{\displaystyle\sum_{k=1}^{\infty} \frac{(\theta\ \beta\ y)^k}{k!\ \Gamma(k)}}. \tag{11}$$

Finally, the conditional expectation becomes

$$E(L^\star \mid X = x) = \frac{\displaystyle\sum_{l=1}^{\infty} \frac{(\theta\ \beta\ y)^l}{(l-1)!\ \Gamma(l)}}{\displaystyle\sum_{k=1}^{\infty} \frac{(\theta\ \beta\ y)^k}{k!\ \Gamma(k)}}. \tag{12}$$

The joint moment generating function (MGF) of X and $L^\star$ can be obtained as

$$\varphi(t,s) = E[e^{tX+sL^\star}] = E[e^{sL^\star} E(e^{tX}|L^\star)] = E\left\{\left[e^s\left(\frac{\beta}{\beta - t}\right)\right]^{L^\star}\right\}, \qquad t < \beta,\ s \in \mathbb{R},$$

which in turn becomes

$$\varphi(t,s) = \frac{1}{e^\theta - 1}\left\{\exp\left[\theta e^s\left(\frac{\beta}{\beta - t}\right)\right] - 1\right\}. \tag{13}$$

With this result, the product moment of X and L^* is derived to be

$$E(XL^*) = \frac{\theta^2\ e^\theta}{\beta(e^\theta - 1)}.$$

Further, it can be shown that for $m, k \geq 0$,

$$E(X^m L^{\star k}) = \frac{\Gamma(m)}{\beta^m(e^\theta - 1)} \sum_{j=1}^{\infty} \frac{j^k \theta^j}{j! B(j,m)},$$

where $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$, for $a, b > 0$.

The MGF of ZTP-EXP distribution $\varphi_X(t)$ can be defined as $\varphi_X(t) = P_{L^\star}(M_T(t))$, where $P_{L^\star}$ is probability generating function (PGF) of zero-truncated Poisson distribution and $M_T(t)$ represents the MGF of exponential distribution. Thus

$$\varphi_X(t) = \frac{1}{e^\theta - 1}\left(\exp\left\{\frac{\theta\beta}{\beta - t}\right\} - 1\right), \quad t < \beta. \tag{14}$$

Hence, the cumulant generating function (CGF) of X is

$$K_X(t) = \ln \varphi_X(t) = \ln\left(\exp\left\{\frac{\theta\beta}{\beta - t}\right\} - 1\right) - \ln\left(e^\theta - 1\right) \tag{15}$$

The mean and variance of X are given, respectively, by

$$E(X) = \frac{\theta}{\beta(1-e^{-\theta})} \quad \text{and} \quad Var(X) = \frac{\theta e^{\theta}}{\beta^2(e^{\theta}-1)^2} \left\{2(e^{\theta}-1) - \theta\right\}.$$

Therefore, the 3-rd moment and skewness measure of X can be expressed, respectively, by

$$E(X-E(X))^3 = \frac{\theta e^{\theta}}{\beta^3(e^{\theta}-1)^3} \left\{\theta e^{\theta}(\theta-6) + 6(\theta + e^{2\theta} - 2e^{\theta} + 1) + \theta^2\right\},$$

and

$$\gamma_1 = \left\{\theta e^{\theta}(2(e^{\theta}-1)-\theta)\right\}^{-3/2} \times \left\{\theta e^{\theta}\left[\theta e^{\theta}(\theta-6) + 6(\theta + e^{2\theta} - 2e^{\theta} + 1) + \theta^2\right]\right\}.$$

The covariance matrix Σ of $(X,\ L^\star)$ can be described as

$$\Sigma = \begin{pmatrix} \frac{\theta e^{\theta}}{\beta^2(e^{\theta}-1)^2}\left[2(e^{\theta}-1)-\theta\right] & \frac{\theta e^{\theta}}{\beta(e^{\theta}-1)^2}\left[e^{\theta}-\theta-1\right] \\ \frac{\theta e^{\theta}}{\beta(e^{\theta}-1)^2}\left[e^{\theta}-\theta-1\right] & \frac{\theta e^{\theta}}{(e^{\theta}-1)^2}\left[e^{\theta}-\theta-1\right] \end{pmatrix}.$$

The correlation coefficient between X and $L^\star$ is given by

$$\rho(\theta) = \left(\frac{e^{\theta}-\theta-1}{2e^{\theta}-\theta-2}\right)^{1/2}.$$

Clearly, the correlation coefficient between X and $L^\star$ depends only on the parametric values of θ. It can be checked that $\rho(\theta)$ increases from 0 to $1/\sqrt{2} \approx 0.71$ when θ increases from 0 to ∞. Therefore, it is clear that the correlation between X and L^* will be always non-negative. It is not very surprising. For any compound distribution, if the baseline distribution has a non-negative support, the correlation becomes non-negative. Moreover, in case of insurance claim it is expected that the total amount paid by the insurance company and the number of claims approved is positively correlated.

2.2 Parameter Estimation Method

Here, we address the problem of computing the MLEs of the two unknown parameters of ZTP-EXP distribution based on a random sample of size n, namely $\{x_1, \ldots, x_n\}$ from ZTP-EXP(β, θ) distribution. The log-likelihood function can be written as

$$l(\beta, \theta) = \sum_{i=1}^{n} \ln f_X(x_i, \beta, \theta). \tag{16}$$

To obtain the MLEs of β and θ, we maximize (16) with respect to the unknown parameters. The normal equations become as

$$\frac{\partial l(\beta,\theta)}{\partial \beta} = 0, \text{ and } \frac{\partial l(\beta,\theta)}{\partial \theta} = 0.$$

It can be easily seen that the maximum likelihood estimates (MLEs) cannot be obtained in explicit forms. We suggest to use EM type algorithm to compute the MLEs of β and θ in this case. Suppose $\{(x_1, l_1), ..., (x_n, l_n)\}$ is a random sample of size n from $(X, L^\star)$. Note that we observe only $\{x_1, ..., x_n\}$ while $\{l_1, ..., l_n\}$ are missing observations. The idea of EM type algorithm is to perform MLEs in the presence of latent variables. It does this by first estimating the values for the latent variables (E-step), optimizing the model (M-step), and then repeating these two steps until convergence. The complete sample log-likelihood function without the additive constant is given by

$$l_c(\beta,\theta) = -\beta\sum_{i=1}^{n} x_i - n\ln(1-e^{-\theta}) + \ln\theta\sum_{i=1}^{n} l_i - n\theta + \ln\beta\sum_{i=1}^{n} l_i + \sum_{i=1}^{n}\Big[(l_i-1)\ln x_i - \ln\Gamma(l_i)\Big].$$

Hence, based on $l_c(\beta,\theta)$, the MLEs of the unknown parameters are given as

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} l_i}{\sum_{i=1}^{n} x_i},$$

and the MLE of θ can be computed by maximizing

$$h(\theta) = \left(\sum_{i=1}^{n} l_i\right)\ln\theta - n\ln(1-e^{-\theta}) - n\theta.$$

The above information provides the basis of the proposed EM type algorithm. Suppose at the j-th stage of the algorithm, the estimates of β and θ are $\beta^{(j)}$ and $\theta^{(j)}$, respectively. For convenience let us write $\Omega^{(j)} = (\beta^{(j)}, \theta^{(j)})$. The E-step and M-step of the EM type algorithm are basically summarized as follows:

E-step: At E-step, we denotes the missing values $l_i^{(j)}$ by $E_i\{\beta^{(j)}, \theta^{(j)}\}$, and it can be obtained by using $\beta^{(j)}$ and $\theta^{(j)}$ as estimates of β and θ, respectively, in (11)

$$E_i\{\beta^{(j)}, \theta^{(j)}\} = \arg\max_{l} P(L^\star = l \mid X = x_i, \Omega^{(j)}).$$

Therefore, the 'pseudo' log-likelihood function becomes

$$\begin{aligned} l_s(\Omega|\Omega^{(j)}) &= -\beta\sum_{i=1}^{n} x_i - n\ln(1-e^{-\theta}) + \ln\theta\sum_{i=1}^{n} E_i\{\beta^{(j)},\theta^{(j)}\} - n\theta + \ln\beta\sum_{i=1}^{n} E_i\{\beta^{(j)},\theta^{(j)}\} \\ &+ \sum_{i=1}^{n}\left[\Big(E_i\{\beta^{(j)},\theta^{(j)}\} - 1\Big)\ln x_i - \ln\Gamma\Big(E_i\{\beta^{(j)},\theta^{(j)}\}\Big)\right]. \end{aligned} \tag{17}$$

M-step: In this step, the 'pseudo' log-likelihood function should be maximized with respect to the unknown parameters. Therefore $\Omega^{(j+1)} = (\beta^{(j+1)},\ \theta^{(j+1)})$ can be obtained as

$$\beta^{(j+1)} = \frac{\sum_{i=1}^{n} E_i\{\beta^{(j)},\theta^{(j)}\}}{\sum_{i=1}^{n} x_i}, \tag{18}$$

and $\theta^{(j+1)}$ is obtained by maximizing the profile 'pseudo' log-likelihood function,

$$\begin{aligned} h(\theta) &= \left(\sum_{i=1}^{n} E_i\{\beta^{(j)}, \theta^{(j)}\}\right) \ln\theta - n\ln(1-e^{-\theta}) - n\theta \\ &= \left(\sum_{i=1}^{n} E_i\{\beta^{(j)}, \theta^{(j)}\}\right) \ln\theta - n\ln(e^{\theta}-1). \end{aligned} \tag{19}$$

It can be shown that the derivative of $h(\theta)$, $h'(\theta)$ changes its sign from positive (+) to negative sign (−) as θ moves from 0 to $+\infty$. As a result of that, for $\sum_{i=1}^{n} E_i\{\beta^{(j)}, \theta^{(j)}\} > n$, $h(\theta)$ has a unique maximum and it can be obtained as the solution of the non-linear equation

$$h^{'}(\theta) = \left(\sum_{i=1}^{n} E_i\{\beta^{(j)}, \theta^{(j)}\}\right)\frac{1}{\theta} - \frac{n\ e^{-\theta}}{1-e^{-\theta}} - n = 0. \tag{20}$$

The used EM type algorithm for computing the unknown parameters for ZTP-EXP distribution is presented in details in Appendix B.

Now we are interested in constructing confidence intervals for the unknown parameters. If $\widehat{\beta}$ and $\widehat{\theta}$ denote the MLEs of β and θ, respectively, then the asymptotic distribution of the MLEs can be obtained as

$$\sqrt{n}(\widehat{\beta}-\beta, \widehat{\theta}-\theta) \xrightarrow{d} N_2(0, J_{obs}^{-1})\ as\ n \to \infty, \tag{21}$$

where $\xrightarrow{d}$ denotes convergence in distribution and J_{obs}^{-1} is the inverse observed Fisher information matrix. For this, using the idea of Louis (1982), the observed Fisher information matrix is obtained from the last step of EM type algorithm. Define the observed Fisher information matrix as follows:

$$J_{obs} = B - SS^T,$$

where, the matrix B denotes the Hessian matrix of the the 'pseudo' log-likelihood function (17) and S is the corresponding gradient vector.

$$B = \begin{pmatrix} \dfrac{\partial^2 l_s}{\partial\beta^2} & \dfrac{\partial^2 l_s}{\partial\beta\ \partial\theta} \\ \\ \dfrac{\partial^2 l_s}{\partial\theta\ \partial\beta} & \dfrac{\partial^2 l_s}{\partial\theta^2} \end{pmatrix} \quad \text{and} \quad S = \left[\frac{\partial l_s}{\partial\beta} \quad \frac{\partial l_s}{\partial\theta}\right]^T.$$

Now we provide the elements of the matrix B and the vector S.

$$\frac{\partial^2 l_s}{\partial\beta^2} = -\frac{1}{\beta^2}\sum_{i=1}^{n} E_i\{\beta^{(j)}, \theta^{(j)}\},\ \frac{\partial^2 l_s}{\partial\beta\ \partial\theta} = 0,$$

$$\frac{\partial^2 l_s}{\partial\theta^2} = -\frac{1}{\theta^2}\sum_{i=1}^{n} E_i\{\beta^{(j)}, \theta^{(j)}\} + \frac{ne^{-\theta}}{(1-e^{-\theta})^2},$$

$$\frac{\partial l_s}{\partial\beta} = \frac{1}{\beta}\sum_{i=1}^{n} E_i\{\beta^{(j)}, \theta^{(j)}\} - \sum_{i=1}^{n} x_i,$$

and

$$\frac{\partial l_s}{\partial \theta} = \frac{1}{\theta}\sum_{i=1}^{n} E_i\{\beta^{(j)}, \theta^{(j)}\} \; - \frac{n\ e^{-\theta}}{1-e^{-\theta}} \; - n.$$

Now, $100(1-\alpha)\%$ confidence intervals of β and θ can be given as

$$(\widehat{\beta} - z_{\alpha/2} f_{11}, \widehat{\beta} + z_{\alpha/2} f_{11}),\ (\widehat{\theta} - z_{\alpha/2} f_{22}, \widehat{\theta} + z_{\alpha/2} f_{22}),$$

where f_{11} and f_{22} are the square root of the diagonal elements of J_{obs}^{-1} and $z_{\alpha/2}$ denotes the α-th percentile point of a standard normal distribution.

3 Bivariate ZTP-EXP Distribution

In this section, the bivariate version of ZTP-EXP is introduced. Also, several properties and theoretical results of the proposed distribution are presented.

3.1 Definitions and Basic Properties

Consider the random variables N, L, L^* and X, same as defined in Section 2. We say,

$$(X,\ V) \stackrel{D}{=} \Big\{(X,\ N)\ |L \geq 1\Big\}, \tag{22}$$

has a bivariate zero truncated Poissin exponential distribution with parameters $(\beta,\ \lambda,\ p)$ and it will be denoted by BZTP-EXP$(\beta,\ \lambda,\ p)$.

Theorem 3.1. If $(X,\ V) \sim$ BZTP-EXP $(\beta,\ \lambda,\ p)$, then the joint CDF of $(X,\ V)$ for $x > 0$ and $v = 1, 2, ...$, is given by

$$P(X \leq x, V \leq v) = \frac{1}{1-e^{-\theta}} \sum_{n=1}^{v}\sum_{l=1}^{n} \frac{\Gamma_{\beta x}(l)}{\Gamma(l)} \binom{n}{l} p^l (1-p)^{n-l}\ \frac{e^{-\lambda}\lambda^n}{n!}, \tag{23}$$

and for $v = 1, 2, ...$

$$P(X \leq x, V = v) = \frac{1}{1-e^{-\theta}} \sum_{l=1}^{v} \frac{\Gamma_{\beta x}(l)}{\Gamma(l)} \binom{v}{l} p^l (1-p)^{v-l}\ \frac{e^{-\lambda}\lambda^v}{v!}. \tag{24}$$

Proof The joint CDF of (X, V) can be written as

$$\begin{aligned}
P(X \leq x, V \leq v) &= \frac{1}{1-e^{-\theta}} \sum_{n=1}^{v} P(X \leq x, N = n, L \geq 1) \\
&= \frac{1}{1-e^{-\theta}} \sum_{n=1}^{v}\sum_{l=1}^{n} P(X \leq y \mid N = n, L = l) P(L = l \mid N = n) P(N = n) \\
&= \frac{1}{1-e^{-\theta}} \sum_{n=1}^{v}\sum_{l=1}^{n} \frac{\Gamma_{\beta x}(l)}{\Gamma(l)} \binom{n}{l} p^l (1-p)^{n-l}\ \frac{e^{-\lambda}\lambda^n}{n!}.
\end{aligned}$$

∎

Theorem 3.2. If $(X,\ V) \sim$ BZTP-EXP $(\beta,\ \lambda,\ p)$, then the marginal CDF of X and PMF of V are

$$F_X(x) = \frac{e^{-\theta}}{1-e^{-\theta}} \sum_{l=1}^{\infty} \frac{\Gamma_{\beta x}(l)}{\Gamma(l)}\ \frac{\theta^l}{l!}, \tag{25}$$

and

$$P(V=v) = \frac{e^{-\lambda}}{v!\ (1-e^{-\lambda})}\ (\lambda^v - (\lambda-\theta)^v), \tag{26}$$

respectively. It is clear from (25) that $Y \sim$ ZTP-EXP$(\beta,\ \theta)$. Specifically, for $p=1$, that is $\theta = \lambda$, we have

$$P(V=v) = \frac{e^{-\lambda}\lambda^v}{v!(1-e^{-\lambda})},$$

Thus, the random variable V has zero-truncated Poisson distribution with parameter λ.

Proof: The CDF of X has already been derived before. The PMF of V can be written as

$$\begin{aligned} P(V=v) &= \frac{1}{1-e^{-\theta}} \sum_{l=1}^{v} \binom{v}{l} p^l (1-p)^{v-l}\ \frac{e^{-\lambda}\lambda^v}{v!} \\ &= \frac{e^{-\lambda}\lambda^v}{v!(1-e^{-\theta})}\ (1-(1-p)^v) = \frac{e^{-\lambda}}{v!(1-e^{-\theta})}\ (\lambda^v - (\lambda-\theta)^v). \end{aligned}$$

■

From Theorems 3.1 and 3.2, it is clear that the parameters (β, λ, p) are identifiable in the joint distribution of (X, V). It can be easily seen from the distribution of V that in V, the parameters λ and θ are identifiable and in the distribution of X, the parameters β and θ are identifiable. Hence, in the joint distribution of (X, V), (β, λ, θ) or equivalently, (β, λ, p) are identifiable.

From Theorem 3.2, we immediately have the conditional CDF and PDF of X given V as follows:

$$F_{X\le x|V=v}(x) = \frac{1}{1-(1-p)^v}\ \sum_{l=1}^{v} \frac{\Gamma_{\beta x}(l)}{\Gamma(l)}\ \binom{v}{l} p^l (1-p)^{v-l}, \tag{27}$$

and

$$f_{X|V=v}(x) = \frac{\beta e^{-\beta x}}{1-(1-p)^v}\ \sum_{l=1}^{v} \frac{(\beta x)^{l-1}}{\Gamma(l)}\ \binom{v}{l} p^l (1-p)^{v-l}. \tag{28}$$

From (28), it can be represented the conditional PDF of X given V can be expressed as follows:

$$X^{\star} \overset{D}{=} \sum_{i=1}^{M} Z_i,$$

where $Z_1, ..., Z_v$ are i.i.d exponential random variable with parameter β and M is zero-truncated binomial random variable with parameters v and p. Also $\{T_i :\ i = 1, 2, ..., v\}$ and M are independent.

Now we derive the joint MGF of X and V for $t < \beta,\ s \in \mathbb{R}$ and for

$$m = \left(1 - p + \frac{p\ \beta}{\beta - t}\right),$$

and it is given by

$$
\begin{aligned}
\varphi_{X,V}(t,s) &= E_{X,V}(e^{tX+sV}) = E_V\left\{e^{sV}E_{Y|V}(e^{tX})\right\} = E_V\left\{e^{sV}\left[\frac{m^V-(1-p)^V}{1-(1-p)^V}\right]\right\}\\
&= \frac{e^{-\lambda}}{1-e^{-\theta}}\sum_{v=1}^{\infty}\frac{1}{v!}\left\{\frac{e^{sv}}{1-(1-p)^v}\left[m^v-(1-p)^v\right]\left[\lambda^v-(\lambda-\theta)^v\right]\right\}.
\end{aligned}
$$

Specifically, for $\lambda=\theta \ \ (p=1)$ and $m=\dfrac{\beta}{\beta-t}$, it reduces

$$
\begin{aligned}
\varphi_{X,V}(t,s) &= \frac{e^{-\theta}}{1-e^{-\theta}}\sum_{v=1}^{\infty}\left\{\frac{\theta^v m^v e^{sv}}{v!}\right\}\\
&= \frac{1}{e^{\theta}-1}\sum_{v=1}^{\infty}\left\{\frac{(\theta m e^s)^v}{v!}\right\} = \frac{e^{\theta m e^s}-1}{e^{\theta}-1}.
\end{aligned}
$$

From the joint MGF of X and V, we can obtain the MGF of V for $t<\beta$, and it is given as

$$
M_V(s) = \frac{e^{-\lambda}}{1-e^{-\theta}}\left[e^{\lambda e^s}-e^{(\lambda-\theta)e^s}\right]. \tag{29}
$$

It follows from (29) that the first two moments of V can be obtained as

$$
E(V)=\lambda+\frac{\theta e^{-\theta}}{1-e^{-\theta}} \text{ and } E(V^2)=\lambda+\lambda^2+\frac{\theta e^{-\theta}(1+2\lambda-\theta)}{1-e^{-\theta}}.
$$

3.2 Statistical Inferences

Now, we present the MLEs of the unknown parameters β, λ and p based on a random sample of size n, namely $\{(x_1,v_1),\ldots,(x_n,v_n)\}$ from BZTP-EXP distribution.

The MLE easily obtained by maximizing the likelihood function with respect to the parameters of our interest. The joint log-likelihood function can be written as

$$
l(\Omega)=\prod_{i=1}^{n} f_{X_i|V_i}(y_i)\ P(V_i=v_i)=\sum_{i=1}^{n}\log f_{X_i|V_i}(x_i)+\sum_{i=1}^{n}\log P(V_i=v_i). \tag{30}
$$

Here, $\Omega=(\beta,\lambda,p)$. Same as before, the MLEs cannot be obtained in explicit forms, it is a 3-dimensional optimization problem. For this, we use also EM type algorithm using the missing information principal. The basic idea is as follows. Suppose $\{(x_1,v_1,l_1),...,(x_n,v_n,l_n)\}$ is a random sample of size n from (X,V,L). But we observe only $\{(x_1,v_1),...,(x_n,v_n)\}$ and $\{l_1,...,l_n\}$ are missing observations. Note that $\{X \mid V=v, L=l\}$ has gamma distribution with shape l and scale β, $\{L=l \mid V=v\}$ has zero-truncated binomial distribution with PDF

$$
P(L=l \mid V=v)=\frac{1}{1-(1-p)^v}\binom{v}{l}p^l(1-p)^{v-l}, \qquad l=1,2,...,v,
$$

V has zero-truncated Poisson distribution with parameter λ. The log-likelihood function without the additive constant based on the complete sample is

$$\begin{aligned} l_c(\Omega) &= -\beta\sum_{i=1}^n x_i + \ln\beta\sum_{i=1}^n l_i + \sum_{i=1}^n (l_i-1)\ln x_i - \sum_{i=1}^n \ln\Gamma(l_i) - \sum_{i=1}^n \ln(1-(1-p)^{v_i}) \\ &+ \ln p\sum_{i=1}^n l_i + \ln(1-p)\sum_{i=1}^n (v_i - l_i) - n\lambda - n\ln(1-e^{-\lambda}) + \ln\lambda\sum_{i=1}^n v_i. \end{aligned} \tag{31}$$

Thus, the MLEs of the unknown parameters are given as

$$\widehat{\beta} = \frac{\sum_{i=1}^n l_i}{\sum_{i=1}^n x_i},$$

and the MLE of λ and p can be obtained by maximizing the two functions h_1 and h_2 respectively, where

$$h_1(\lambda) = \left(\sum_{i=1}^n v_i\right)\ln\lambda - n\ln(1-e^{-\lambda}) - n\lambda,$$

and

$$h_2(p) = \ln p\sum_{i=1}^n l_i + \ln(1-p)\sum_{i=1}^n (v_i - l_i) - \sum_{i=1}^n \ln(1-(1-p)^{v_i}).$$

Now, the EM type algorithm can be obtained as follows. Suppose at the j-th stage of the algorithm, $\Omega^{(j)} = (\beta^{(j)}, \lambda^{(j)}, p^{(j)})$ denote the estimates of the unknown parameters. The E-step and M-step of the EM type algorithm are basically exemplified in as follows.

E-step: The E-step is the same as before. Hence, the 'pseudo' log-likelihood function at the j-th stage can be written as follows:

$$\begin{aligned} l_s(\Omega \mid \Omega^{(j)}) &= -\beta\sum_{i=1}^n x_i + \ln\beta\sum_{i=1}^n l_i^{(j)} + \sum_{i=1}^n (l_i^{(j)}-1)\ln x_i - \sum_{i=1}^n \ln\Gamma(l_i^{(j)}) \\ &- \sum_{i=1}^n \ln(1-(1-p)^{v_i}) + \ln p\sum_{i=1}^n l_i^{(j)} + \ln(1-p)\sum_{i=1}^n (v_i - l_i^{(j)}) \\ &- n\lambda - n\ln(1-e^{-\lambda}) + \ln\lambda\sum_{i=1}^n v_i. \end{aligned} \tag{32}$$

Note that

$$l_i^{(j)} = \arg\max_l P(L = l \mid V = v_i, p^{(j)}). \tag{33}$$

M-step: M-step can be obtain by maximizing the 'pseudo' log-likelihood function with respect the unknown parameters, therefore $\Omega^{(j+1)} = (\beta^{(j+1)},\ \lambda^{(j+1)},\ p^{(j+1)})$ can be obtained as

$$\beta^{(j+1)} = \frac{\sum_{i=1}^n l_i^{(j)}}{\sum_{i=1}^n x_i}. \tag{34}$$

Moreover, $\lambda^{(j+1)}$ and $p^{(j+1)}$ can be obtained respectively as the solution of the non-linear two equations

$$h_1^{'}(\lambda) = \left(\sum_{i=1}^{n} v_i\right)\frac{1}{\lambda} - \frac{n\ e^{-\lambda}}{1-e^{-\lambda}} - n = 0, \tag{35}$$

and

$$h_2^{'}(p) = \left(\sum_{i=1}^{n} l_i^{(j)}\right)\frac{1}{p} - \frac{1}{1-p}\left(\sum_{i=1}^{n}(v_i - l_i^{(j)})\right) - \sum_{i=1}^{n}\left[\frac{v_i(1-p)^{v_i-1}}{1-(1-p)^{v_i}}\right] = 0. \tag{36}$$

The steps of the EM type algorithm for computing the unknown parameters for BZTP-EXP distribution are presented in Appendix C.

Now, at the last step of the EM type algorithm, by method of Louis (1982) we can obtain approximate confidence intervals of the unknown parameters. If $\widehat{\beta},\ \widehat{\lambda}$ and $\widehat{p}$ are the MLEs of of $\beta,\ \lambda$ and p respectively, under the usual regularity conditions, the asymptotic distribution of the MLEs can be obtained as

$$\sqrt{n}(\widehat{\beta}-\beta,\ \widehat{\lambda}-\lambda,\ \widehat{p}-p) \xrightarrow{d} N_3(0, J_{obs}^{-1}) \quad as \quad n \to \infty.$$

For this, we provide the observed information matrix which can be used to obtain the approximate confidence intervals of the unknown parameters. The observed Fisher information matrix can be written in as $J_{obs} = B - SS^T$. Here, the matrix B denotes the Hessian matrix of the the 'pseudo' log-likelihood function (32) and S is the corresponding gradient vector.

$$B = \begin{pmatrix} \dfrac{\partial^2 l_s}{\partial\ \beta^2} & \dfrac{\partial^2 l_s}{\partial\beta\ \partial\lambda} & \dfrac{\partial^2 l_s}{\partial\beta\ \partial p} \\ \dfrac{\partial^2 l_s}{\partial\lambda\ \partial\beta} & \dfrac{\partial^2 l_s}{\partial\lambda^2} & \dfrac{\partial^2 l_s}{\partial\lambda\ \partial p} \\ \dfrac{\partial^2 l_s}{\partial p\ \partial\beta} & \dfrac{\partial^2 l_s}{\partial p\ \partial\lambda} & \dfrac{\partial^2 l_s}{\partial p^2} \end{pmatrix} \quad \text{and} \quad S = \left[\frac{\partial l_s}{\partial\beta} \quad \frac{\partial l_s}{\partial\lambda} \quad \frac{\partial l_s}{\partial p}\right]^T.$$

Now we provide the elements of the matrix B and the vector S.

$$\frac{\partial^2 l_s}{\partial\beta^2} = -\frac{1}{\beta^2}\sum_{i=1}^{n} l_i^{(j)}, \qquad \frac{\partial^2 l_s}{\partial\beta\ \partial\lambda} = 0, \qquad \frac{\partial^2 l_s}{\partial\beta\ \partial p} = 0,$$

$$\frac{\partial^2 l_s}{\partial\lambda^2} = -\frac{1}{\lambda^2}\sum_{i=1}^{n} v_i + \frac{ne^{-\lambda}}{(1-e^{-\lambda})^2}, \qquad \frac{\partial^2 l_s}{\partial\lambda\ \partial p} = 0,$$

$$\frac{\partial^2 l_s}{\partial p^2} = -\frac{1}{p^2}\sum_{i=1}^{n} l_i^{(j)} - \frac{1}{(1-p)^2}\left(\sum_{i=1}^{n}(v_i - l_i^{(j)})\right) + \sum_{i=1}^{n}\left[\frac{v_i(v_i-1)(1-p)^{v_i-2} + (1-p)^{2v_i-2}}{(1-(1-p)^{v_i})^2}\right],$$

$$\frac{\partial l_s}{\partial\beta} = -\sum_{i=1}^{n} x_i + \frac{1}{\beta}\sum_{i=1}^{n} l_i^{(j)}, \ \frac{\partial l_s}{\partial\lambda} = \frac{1}{\lambda}\left(\sum_{i=1}^{n} v_i\right) - \frac{n\ e^{-\lambda}}{1-e^{-\lambda}} - n,$$

and

$$\frac{\partial l_s}{\partial p} = \frac{1}{p}\sum_{i=1}^{n} l_i^{(j)} - \frac{1}{1-p}\sum_{i=1}^{n}(v_i - l_i^{(j)}) - \sum_{i=1}^{n}\left[\frac{v_i(1-p)^{v_i-1}}{1-(1-p)^{v_i}}\right].$$

Therefore, $100(1-\alpha)\%$ confidence intervals of β, λ and p can be obtained as

$$(\widehat{\beta} - z_{\alpha/2} f_{11}, \widehat{\beta} + z_{\alpha/2} f_{11}), \qquad (\widehat{\lambda} - z_{\alpha/2} f_{22}, \widehat{\lambda} + z_{\alpha/2} f_{22}), \qquad (\widehat{p} - z_{\alpha/2} f_{33}, \widehat{p} + z_{\alpha/2} f_{33}),$$

respectively. Here f_{11}, f_{22} and f_{33} are the square root of the diagonal elements of J_{obs}^{-1} and $z_{\alpha/2}$ denotes the α-th percentile point of a standard normal distribution.

4 Numerical Results

In this section, we present the results of a simulation study designed to study the performance of the proposed EM type algorithm in computing the MLEs for ZTP-EXP distribution and BZTP-EXP distribution using different parameters values and sample sizes. Also, analysis of the Swedish buss insurance data set is presented for illustrative purposes. All computations are performed using R Package.

4.1 Results for Simulation Study

In this subsection, we present some results for the ZTP-EXP and BZTP-EXP distributions based on Monte Carlo (MC) simulations to compare the performance of the proposed models for different sample sizes, and for different parameter values. For a given set of parameter values for β and θ, we generate a random sample of size n from the ZTP-EXP distribution as we discussed earlier in Section 2. The simulation study is conducted for the parametric values $\beta = 1$ and $\theta = 0.005, 0.5, 1.0, 1.5, 2, 2.5$ and simple sizes $n = 25, 50, 100, 150, 200, 250, 300, 350, 400$. Then, for the given generated sample the proposed EM type algorithm is used to estimate the MLEs of the unknown parameters β and θ. The EM type algorithm is stopped when the convergence is reached (the absolute difference between successive estimates are less than 10^{-5} for the two estimates). The generation process is repeated 1000 times and then the average estimates (AEs) and the mean squared errors (MSEs) of the estimates of β and θ are computed. It can be seen from Tables 1 and 2 that as the sample size increases the AEs and associated MSEs decrease. This implies that the proposed EM type algorithm is working well for the considered distribution.

It is worth mentioning that the estimates of θ in Table 2 for the case when $\theta = 0.005$ are almost the true value and the resulting MSEs diminish to 0 due to the following fact. As θ tends to be 0, the PMF of L^* converges to 1 if $L^* = 1$ and 0 when $L^* > 1$. This implies that the joint PDF of X and L^* is $EXP(\beta)$ for $L^* = 1$ and 0, elsewhere. For this, the MLE of the unknown parameter β is computed explicitly as $\hat{\beta} = 1/\bar{X}$.

For estimating the unknown parameters involved in the BZTP-EXP distribution, we use the same generation process discussed earlier. A random sample of size n is generated from BZTP-EXP distribution by considering a given set of three parameter values for β, λ and p. Also, the generated random sample is used to utilize the EM type algorithm to estimate the MLEs of the unknown parameters β, λ and p. For each case, the EM type algorithm is stopped when the absolute difference between successive estimates are less than 10^{-5} for the three estimates. The process is repeated 1000 times based on the parametric values: $\beta = 1$, and $(\lambda, p) = (3, 0.5), (3.5, 0.5), (3, 0.6)$ with sample sizes $n = 25, 50, 100, 150, 200, 250, 300, 350, 400$. The results of AEs and MSEs are summarized in

Table 3. In this case also, the AEs and the MSEs decrease as the sample size increases. It indicates that the proposed EM type algorithm is working well. It verifies the consistency property of the MLEs.

Table 1: The AE and the associated MSEs of the MLEs for the ZTP-EXP distribution when $\beta = 1.0$, $\theta = 0.5$, 1.0 and 1.5.

n	Par	$\beta = 1.0, \theta = 0.5$		$\beta = 1.0, \theta = 1.0$		$\beta = 1.0, \theta = 1.5$	
		AE	MSE	AE	MSE	AE	MSE
25	β	1.2978	0.3128	1.2399	0.3234	1.3634	0.4484
	θ	0.9952	0.7974	1.3827	1.0110	2.2738	1.9306
50	β	1.1553	0.1027	1.1795	0.1320	1.1360	0.1402
	θ	0.7703	0.3666	1.3630	0.5504	1.7973	0.7137
100	β	1.1376	0.0988	1.1424	0.0854	1.1004	0.0766
	θ	0.7481	0.3284	1.3060	0.4130	1.7303	0.4945
150	β	1.0944	0.0892	1.0785	0.0651	1.0942	0.0724
	θ	0.6714	0.2516	1.1626	0.3047	1.7080	0.4336
200	β	1.0979	0.0424	1.0682	0.0464	1.0410	0.0348
	θ	0.6421	0.1449	1.1238	0.2342	1.6134	0.2488
250	β	1.0660	0.0360	1.0510	0.0368	1.0599	0.0339
	θ	0.6375	0.1277	1.1138	0.1774	1.6041	0.1904
300	β	1.0599	0.0237	1.0469	0.0338	1.0374	0.0310
	θ	0.6188	0.0979	1.0686	0.1552	1.5533	0.1859
350	β	1.0361	0.0261	1.0151	0.0233	1.0313	0.0207
	θ	0.5639	0.0719	1.0162	0.1198	1.5424	0.1262
400	β	1.0235	0.0110	1.0117	0.0162	1.0197	0.0131
	θ	0.5361	0.0391	1.0023	0.0811	1.5393	0.0830

Table 2: The AE and the associated MSEs of the MLEs for the ZTP-EXP distribution when $\beta = 1.0$, $\theta = 0.005, 2.0$ and 2.5.

n	Par	$\beta = 1.0, \theta = 0.005$		$\beta = 1.0, \theta = 2.0$		$\beta = 1.0, \theta = 2.5$	
		AE	MSE	AE	MSE	AE	MSE
25	β	1.0234	0.0516	1.1601	0.2015	1.1492	0.1924
	θ	0.0051	1.09×10^{-06}	2.3915	1.6809	2.8894	1.6378
50	β	1.0126	0.0230	1.1307	0.1248	1.1032	0.0675
	θ	0.0050	5.29×10^{-07}	2.2878	0.8269	2.8059	0.8134
100	β	1.0114	0.0118	1.1183	0.0664	1.0978	0.0630
	θ	0.0049	2.71×10^{-07}	2.2911	0.5587	2.6829	0.5585
150	β	1.0168	0.0064	1.0936	0.0435	1.0580	0.0345
	θ	0.0049	1.45×10^{-07}	2.2711	0.3750	2.6472	0.2766
200	β	1.0092	0.0047	1.0386	0.0265	1.0445	0.0212
	θ	0.0049	1.13×10^{-07}	2.0927	0.2244	2.6329	0.2398
250	β	1.0038	0.0042	1.0328	0.0222	1.0306	0.0238
	θ	0.0050	1.03×10^{-07}	2.0771	0.1788	2.5516	0.2395
300	β	1.0074	0.0033	1.0262	0.0136	1.0245	0.0181
	θ	0.0049	8.15×10^{-08}	2.0425	0.1111	2.5488	0.1797
350	β	1.0035	0.0030	1.0209	0.0143	1.0182	0.0174
	θ	0.0049	7.51×10^{-08}	2.0526	0.0986	2.5409	0.1640
400	β	0.9960	0.0024	1.0137	0.0112	1.0111	0.0100
	θ	0.0050	6.20×10^{-08}	2.0439	0.0929	2.5169	0.1025

Table 3: The AE and the associated MSEs of the MLEs for the BZTP-EXP distribution when $\beta = 1.0$, $\lambda = 3$, 3.5, $p = 0.5$ and 0.6.

n	Par	$\beta = 1.0,\ \lambda = 3,\ p = 0.5$		$\beta = 1.0,\ \lambda = 3.5,\ p = 0.5$		$\beta = 1.0,\ \lambda = 3,\ p = 0.6$	
		AE	MSE	AE	MSE	AE	MSE
25	β	0.8503	0.1242	0.9440	0.0837	0.8036	0.1129
	λ	2.9263	0.1447	3.4429	0.1639	2.9786	0.1379
	p	0.3006	0.1035	0.4314	0.0438	0.3474	0.1254
50	β	0.9178	0.0754	0.9801	0.0426	0.8155	0.0876
	λ	2.9839	0.0559	3.4666	0.0680	2.9852	0.0649
	p	0.3841	0.0626	0.4795	0.0196	0.3853	0.0957
100	β	0.9322	0.0490	1.0093	0.0231	0.9579	0.0355
	λ	2.9786	0.0323	3.4744	0.0501	2.9881	0.0278
	p	0.4273	0.0338	0.4946	0.0124	0.5402	0.0375
150	β	0.9402	0.0391	1.0199	0.0200	0.9922	0.0235
	λ	2.9993	0.0226	3.4820	0.0264	3.0095	0.0213
	p	0.4469	0.0257	0.4973	0.0099	0.5680	0.0226
200	β	0.9477	0.0345	1.0248	0.0116	0.9923	0.0161
	λ	3.0013	0.0160	3.5141	0.0224	3.0056	0.0188
	p	0.4523	0.0225	0.5069	0.0051	0.5788	0.0153
250	β	0.9713	0.0259	1.0282	0.0050	1.0103	0.0116
	λ	3.0108	0.0111	3.5281	0.0116	3.0077	0.0121
	p	0.4650	0.0177	0.5196	0.0004	0.5945	0.0080
300	β	0.9740	0.0258	1.0316	0.0033	1.0072	0.0089
	λ	2.9936	0.0109	3.5298	0.0112	3.0189	0.0102
	p	0.4681	0.0161	0.5171	0.0003	0.5933	0.0080
350	β	1.0001	0.0133	1.0142	0.0026	1.0045	0.0077
	λ	3.0222	0.0105	3.4907	0.0098	2.9996	0.0090
	p	0.4896	0.0081	0.5187	0.0003	0.5972	0.0045
400	β	0.9910	0.0113	1.0217	0.0023	1.0123	0.0058
	λ	3.0125	0.0076	3.5036	0.0080	3.0026	0.0082
	p	0.4932	0.0065	0.5166	0.0002	0.6026	0.0040

4.2 Data Analysis and Discussion:

In this subsection, we discuss the analysis of real life data representing Swedish motor insurance data set. This data set is collected from seven geographical zones in 1977 by the Swedish committee on the analysis of risk premium in motor insurance and it is available in Andrews and Herzberg [1] and in the R package CASdatasets [12]. The considered data contains the total amount paid by the insurance company (X) and the total number of claims received by the insurance company (V). For computational purposes, each observation of total payments amount is divided by 100 for computational purposes, it is not going to affect in the inference issues.

The histogram with the fitted PDF of the above data set is presented in Figure 2. Some basic descriptive statistics of the total amount paid by the insurance company are summarized in Table 4. The coefficient skewness indicates that the data are positively skewed and the Figure 2 supports this point.

Table 4: Basic descriptive statistics of Y.

Mean	Median	Std.Dev	Q_1	Q_3	skewness
32.966	21.706	28.076	12.648	44.966	1.685

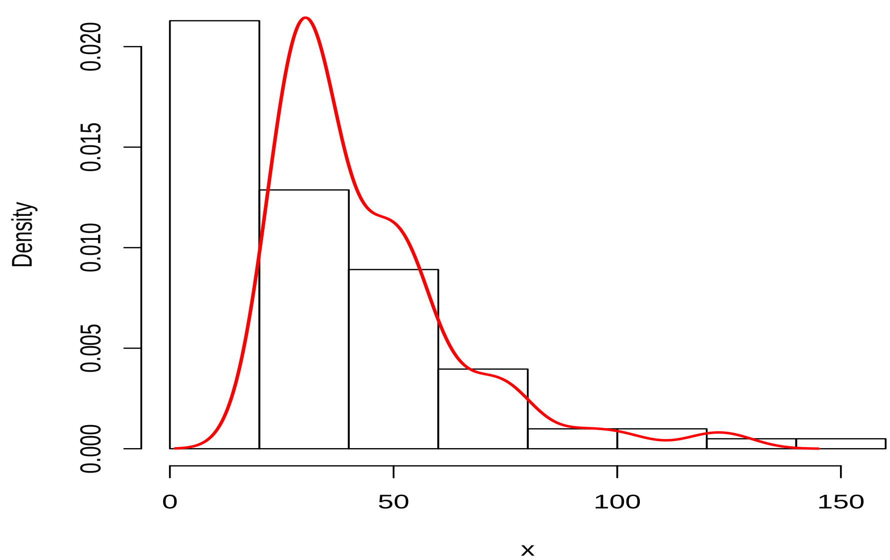

Figure 2: Histogram and fitted PDF for the total amount paid by insurance company.

Now, we would like to fit BZTP-EXP distribution to the bivariate payment data set and use the EM type algorithm for obtaining the MLEs of the unknown parameters. With the starting values $(\beta^{(0)}, \lambda^{(0)}, p^{(0)}) = (0.1227, 4.4280, 0.5)$, it took the EM type algorithm 3 iterations to converge. The resulting MLEs of the unknown parameters and associated log-likelihood (ll) are

$$\hat{\beta}^{(3)} = 0.0964, \quad \hat{\lambda}^{(3)} = 5.7539, \quad \hat{p}^{(3)} = 0.5376, \quad \text{ll} = -400.374.$$

The corresponding 95% confidence intervals of β, λ and p respectively are

$$(0.0858, 0.1069), \quad (5.2825, 5.2253), \quad (0.4954, 0.5799).$$

To verify whether the estimates obtained using the proposed EM type algorithm actually converge to the MLEs, we have performed extensive grid search method with a grid size 0.001 for β, λ and p, and it is observed that they match. Further, we have used different other initial estimates, for example $(\beta^{(0)}, \lambda^{(0)}, p^{(0)}) = (0.2087, 3.0840, 0.55)$, the EM type algorithm converges to the same result. This ensures that the proposed EM type algorithm actually converges to the MLEs and the initial estimates do not affect the convergence of the EM type algorithm.

Further, the ZTP-EXP is fitted to the univariate data set. The EM type algorithm have been used to compute the MLEs of unknown parameters. The initial values for the unknown parameters are $\beta^{(0)} = 0.1209$ and $\theta^{(0)} = 2.1706$. The EM type algorithm stops after twelve iterations and the MLEs of unknown parameters and the corresponding log-likelihood (ll) are

$$\hat{\beta}^{(12)} = 0.0901, \quad \hat{\theta}^{(12)} = 2.7884, \quad \text{ll} = -412.973.$$

The associated 95% confidence bounds of β and θ are, respectively, presented below:

$$(0.0799, 0.1003), \quad (2.4395, 3.1374).$$

The bivariate compound negative binomial exponential (BCNB-EXP) distribution is fitted to the bivariate data set. For more details on the model, one may refer to the derivations presented in Appendix A. Based on the initial strategies given above with starting values of $(\beta^{(0)}, p^{(0)}, \nu^{(0)}, r^{(0)}) = (0.1227, 0.6, 0.3508, 3.450)$, we readily obtain the MLEs of β, p, ν and r after eleven iterations as well as the corresponding log-likelihood (ll) as

$$\hat{\beta}^{(11)} = 0.1036, \quad \hat{p}^{(11)} = 0.5821, \quad \hat{\nu}^{(11)} = 0.3985, \quad \hat{r}^{(11)} = 3.4551, \quad \text{ll} = -442.264.$$

The corresponding 95% confidence intervals of β, p, ν and r, respectively, are

$$(0.0926, 0.1145), \quad (0.5407, 0.6236), \quad (0.3587, 0.4638), \quad (3.3110, 3.5992).$$

The compound negative binomial exponential (CNB-EXP) distribution is fitted to the univariate data set. By using the initial values of $(\beta^{(0)}, p^{(0)}, r^{(0)})$ as $(0.3296, 0.180, 2.0)$, the MLEs of β, p and r and the corresponding log-likelihood (ll) are reached after three iterations. These estimates are computed to be

$$\hat{\beta}^{(3)} = 0.3009, \quad \hat{\eta}^{(3)} = 0.2014, \quad \hat{r}^{(3)} = 2.0011, \quad \text{ll} = -450.146.$$

The corresponding 95% confidence intervals of β, η and r, respectively, are

$$(0.2823, 0.3195), \quad (0.1766, 0.2262), \quad (1.7935, 2.2088).$$

The compound geometric exponential (CG-EXP) distribution is fitted to the univariate data set (see, [15]). The EM type algorithm is used to compute the MLEs of the unknown parameters. We started it with initial values of $\beta^{(0)}$ and $p^{(0)}$ as 0.2807 and 0.5, respectively. The MLEs of unknown parameters are obtained after eight iterations and the finale estimate of β and p and the corresponding log-likelihood (ll) are

$$\hat{\beta}^{(8)} = 0.0507, \quad \hat{p}^{(8)} = 0.5976, \quad \text{ll} = -524.259.$$

Table 5: The goodness of fit tests for insurance data set

Distribution	AIC	AICc	HQIC	BIC	K-S	p-value
BZTP-EXP	806.748	806.995	809.924	814.593	0.0956	0.3708
ZTP-EXP	829.946	830.068	836.063	839.176	0.1065	0.3106
BCNB-EXP	892.528	892.944	896.762	902.988	0.1146	0.1789
CNB-EXP	906.292	906.539	909.468	914.137	0.1278	0.0994
CG-EXP	1052.518	1052.640	1054.635	1057.748	0.1466	0.0385
EXP	916.080	910.120	911.138	912.695	0.1396	0.0557
EXP-PAR	900.766	901.182	905.000	911.226	0.1149	0.1493
GAM-LN	811.550	812.185	816.843	824.625	0.0765	0.5949

The corresponding 95% confidence intervals of β and p, respectively, are

$$(0.0431, 0.0584), \quad (0.5237, 0.6715).$$

As mentioned in Section 2, the exponential (EXP) distribution is a special case of ZTP-EXP when θ tends to be 0. The exponential distribution is fitted it to univariate data set. The MLE of β, the corresponding log-likelihood (ll) and their corresponding 95% confidence interval (within bracket) are obtained as

$$\hat{\beta} = 0.0303 \ (\mp 0.0059), \ \ \text{ll} = -454.048.$$

The natural question that arises, what is the most appropriate model for fitting the data set?. To check that, we compute Akaike Information Criterion (AIC), Akaike Information Criterion correction (AICc), Hannan-Quinn Information Criterion (HQIC), Bayesian Information Criterion (BIC) and Kolmogorov-Smirnov (K-S) statistic. Table 5 represents the obtained results. According to the statistics of this table. Since values of AIC, AICc, HQIC, BIC and K-S for BZTP-EXP are smaller than ZTP-EXP, BCNB-EXP, CNB-EXP, CG-EXP and EXP distributions, the proposed BZTP-EXP is indeed an appropriate distribution for the data set. The plots of the empirical CDF and fitted CDF in Figure 3 ensures this conclusion. Figure 4 displays QQ-plot of the fitted distributions on the data set. Other alternative models for analyzing heavy-tailed and skewed insurance loss data are finite mixture models of univariate distributions (see, Miljkovic and Grun [22]). Under the parametric setting in Miljkovic and Grun [22], two mixture models are also considered for comparison purposes, namely, Exponential-Pareto (EXP-PAR) and Gamma-Lognormal (GAM-LN). It is seen from Table 5 that BZTP-EXP performs better comparing with EXP-PAR while the GAM-LN model outperforms BZTP-EXP model. This is expected as the mixing and compounding models provide flexible tools for analyzing the loss data.

It is worth mentioning that the dependence structure between the total amount paid by the insurance company, X and the size of claims, L^* have been computed via the estimated correlation coefficient, and it is computed to be $\rho(\theta) = 0.6706$. This in turns out that X and L^* have a moderate positive relationship.

5 Tail Risk Measures

In order to evaluate exposure to risk, different risk measures and their properties have been widely studied in the literature (see Grun et al. [13], Tomarchio and Punzo [27], McNeil et al. [21], and references therein). In this section, we calculate two well-known risk measures for the proposed models namely: Value at Risk (VaR) and Tail value at Risk (TVaR). These measures play a crucial role in portfolio optimization under uncertainty. The VaR, known as the quantile risk measure is always specified with a given level of significance say q. The VaR of a random variable X is the q^{th} quantile of its CDF, and it is defined as

$$\mathrm{VaR}_q(X) = \inf\{x : F(x) > q\} = F_X^{-1}(q), \;\; 0 < q < 1,$$

where F is the CDF of X. In this line, the historical value at risk (HR VaR) is the most common non-parametric approach for obtaining the VaR measure. It is calculated by taking a percentile of the returns r_t and multiplying it by the total value of the loss portfolio and square root of holding period M, the time period over which the losses may occur. Here M is equal to 1 and

$$r_t = \log\left(\frac{S_t - S_{t-1}}{S_{t-1}}\right),$$

with S_t being the value of the loss portfolio at time t.

Another related measure is the TVaR known as conditional tail expectation (CTE) or the expected shortfall (ES). It quantifies the average of losses above the VaR for some given confidence level, that is,

$$\mathrm{TVaR}_q(X) = E[X \mid X > \mathrm{VaR}_q(X)] = \frac{1}{1-q}\int_{\mathrm{VaR}_q}^{\infty} x f(x)\, dx.$$

Note that, a model with higher values of the risk measures (VaR and TVaR) is said to have heavier tail. For assessing how the ZTP-EXP and BZTP-EXP models behave with respect to these measures, a simulation study is performed and its results on VaR and TVaR are summarized in Tables 6-7, and graphically displayed in Figures 5-6. It can be seen that the BZTP-EXP model has higher values of the risk measures than that of the ZTP-EXP model. It is evident that Figures 5 and 6 ensure the superiority of BZTP-EXP to ZTP-EXP.

Moreover, VaR, HR VaR and TVaR measures of the fitted models using the estimated values of the parameters for the insurance claim data set are computed and reported in Tables 8 and 9. It is observed from these tables that the VaR, HR VaR and TVaR values of BZTP-EXP and GAM-LN models are close to their corresponding empirical values at the different significance levels. This shows that BZTP-EXP and GAM-LN can be considered as good competing models for fitting the insurance claim data sets.

Table 6: Simulation results of VaR and TVaR for n = 100.

Distribution	Parameters	Level of significance	VaR	TVaR
ZTP-EXP	β =1	0.700	1.7358	3.3112
	θ =1	0.750	2.1018	3.5907
		0.800	2.4316	3.8602
		0.850	2.8532	3.9179
		0.900	3.0927	4.9017
		0.950	5.4823	6.4388
		0.999	6.6428	7.3170
BZTP-EXP	β =1	0.700	2.6515	4.3759
	λ =3.5	0.750	2.7175	4.4392
	p =0.5	0.800	3.5571	5.5350
		0.850	3.7341	5.9837
		0.900	5.2052	7.1118
		0.950	6.3541	8.0408
		0.999	9.5089	11.3052

Table 7: Simulation results of VaR and TVaR for n = 150.

Distribution	Parameters	Level of significance	VaR	TVaR
ZTP-EXP	β =1	0.700	2.3857	4.0128
	θ =1.5	0.750	2.6666	4.5407
		0.800	3.2407	4.7694
		0.850	3.4611	5.4945
		0.900	4.1153	5.9498
		0.950	5.1994	7.1971
		0.999	6.9865	8.1566
BZTP-EXP	β =1	0.700	2.4852	4.1516
	λ =3	0.750	3.1396	4.8872
	p =0.65	0.800	3.3783	5.4230
		0.850	3.7165	5.9698
		0.900	5.3401	6.9227
		0.950	6.1815	7.6923
		0.999	9.3339	10.7735

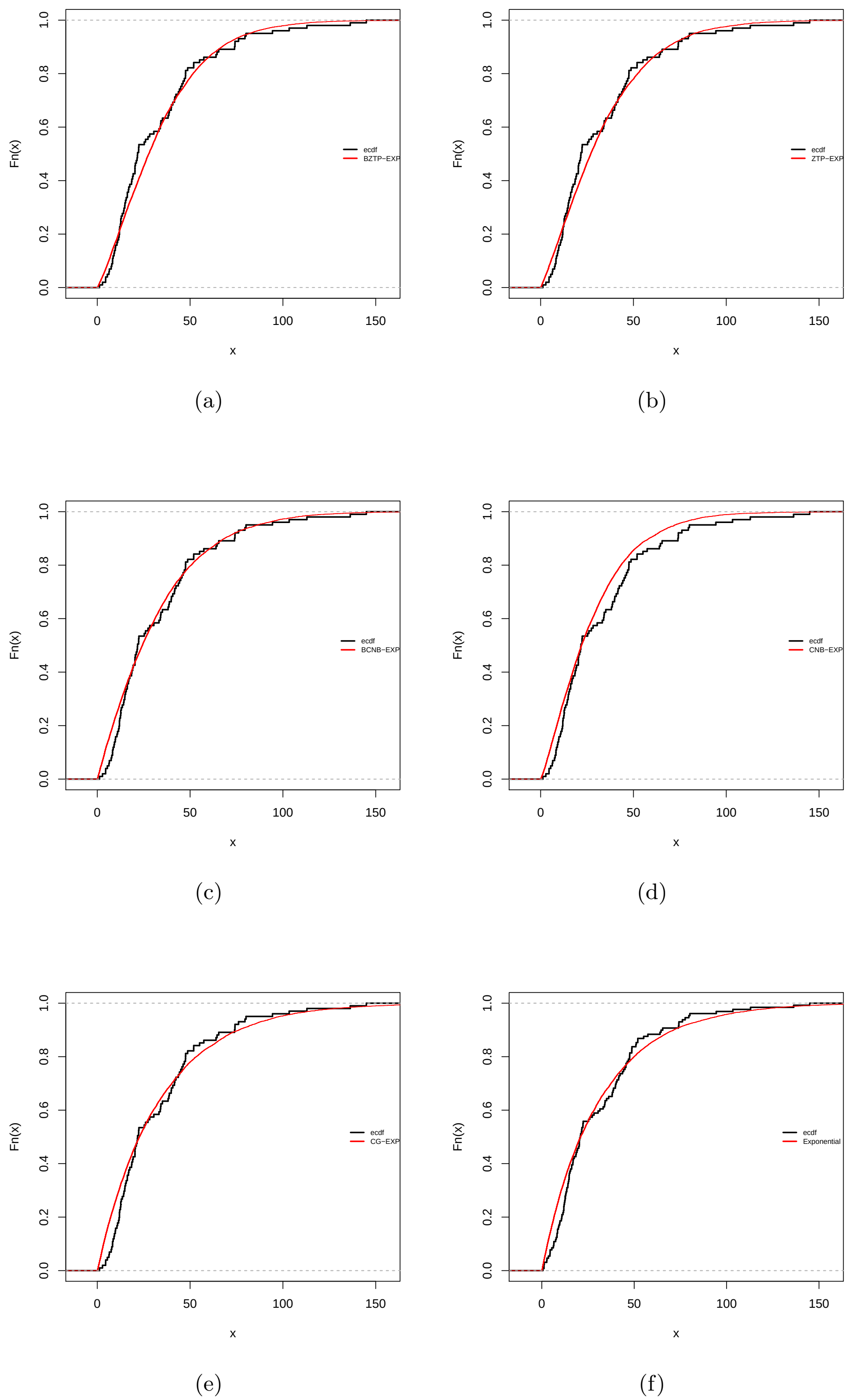

Figure 3: The ECDF and fitted CDFs for different distributions: (a) BZTP-EXP (b) ZTP-EXP (c) BCNB-EXP (d) CNB-EXP (e) CG-EXP (f) Exponential.

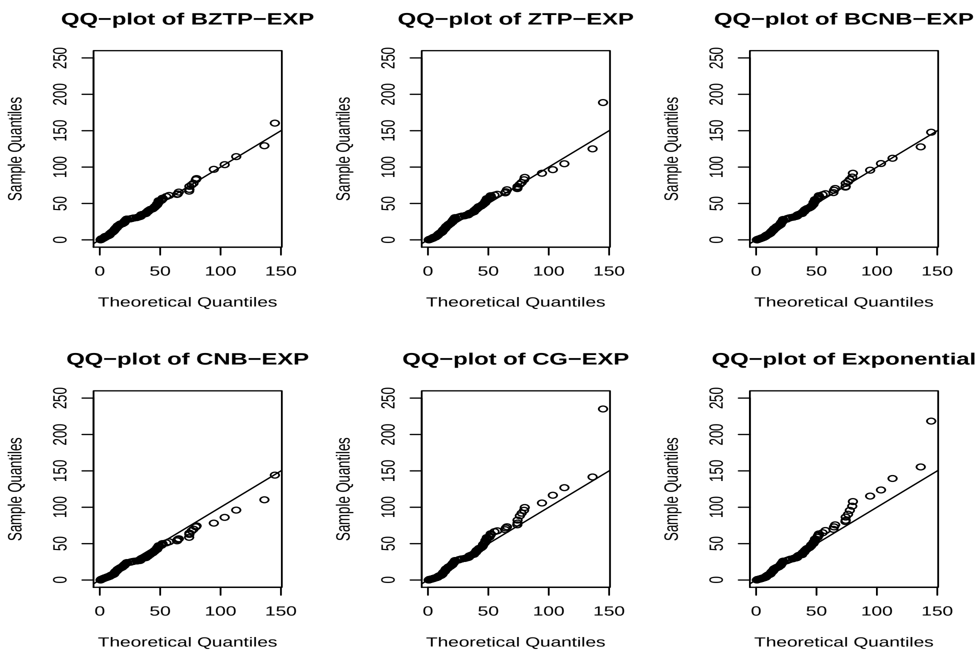

Figure 4: QQ plots of BZTP-EXP, ZTP-EXP, BCNB-EXP, CNB-EXP and EXP models for the insurance data set.

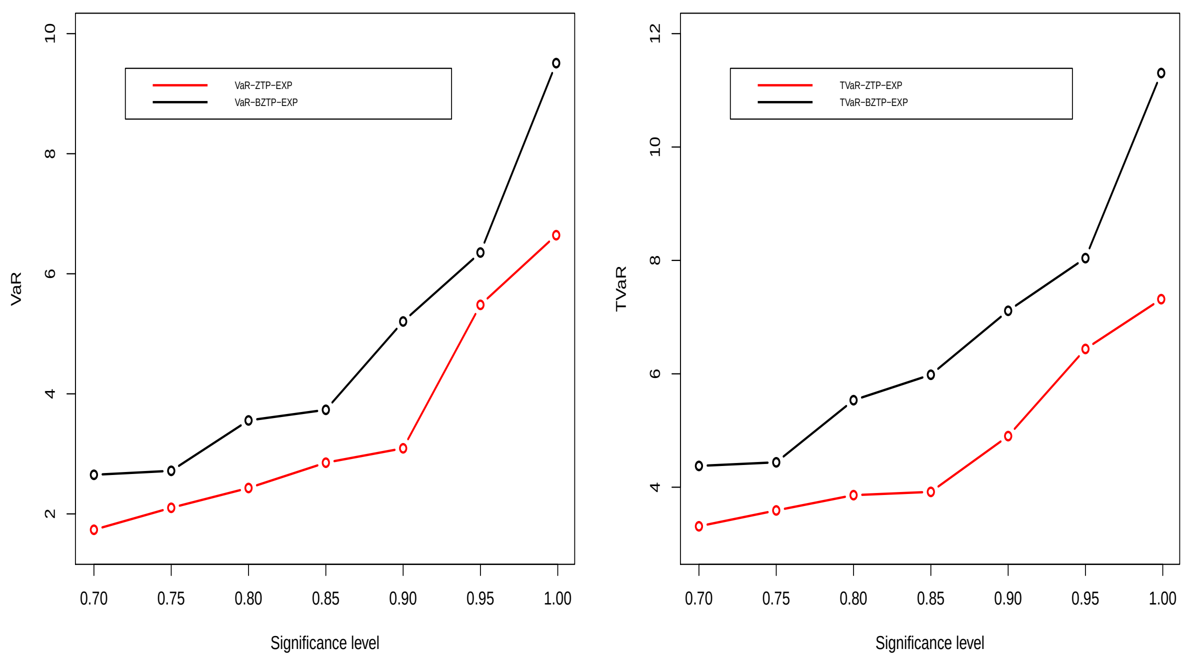

Figure 5: Plots for the value at risk (VaR) and tail value at risk (TVaR) of the ZTP-EXP and BZTP-EXP models for n =100.

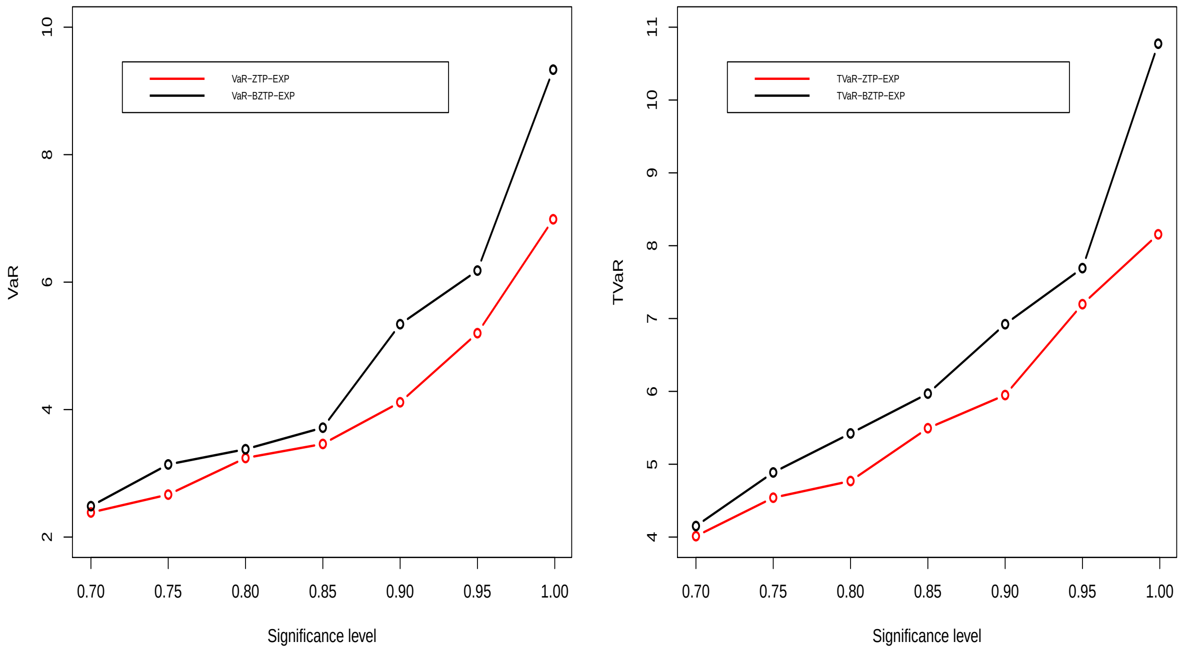

Figure 6: Plots for the value at risk (VaR) and tail value at risk (TVaR) of the ZTP-EXP and BZTP-EXP models for n =150.

Table 8: Results of VaR, HR VAR and TVaR at the 0.70, 0.75 and 0.85 significance levels using the insurance claim data.

Distribution	$\text{VaR}_{0.70}$	HR $\text{VaR}_{0.70}$	$\text{TVaR}_{0.70}$	$\text{VaR}_{0.75}$	HR $\text{VaR}_{0.75}$	$\text{TVaR}_{0.75}$	$\text{VaR}_{0.85}$	HR $\text{VaR}_{0.85}$	$\text{TVaR}_{0.85}$
Empirical	41.564	40.279	66.917	44.966	43.938	71.619	55.124	52.753	86.759
BZTP-EXP	42.497	42.422	64.561	46.916	45.682	75.308	57.543	58.336	81.653
ZTP-EXP	40.415	38.428	62.416	42.936	41.755	74.543	50.304	51.296	74.225
CNB-EXP	34.638	34.264	60.311	39.084	38.371	67.992	50.273	48.130	73.416
NB-EXP	34.346	32.134	57.847	38.628	36.267	68.779	50.599	46.593	71.476
CG-EXP	32.579	31.591	43.952	37.022	35.215	50.884	44.599	42.317	60.602
EXP	33.308	32.350	44.948	37.457	35.821	53.054	47.019	44.360	65.870
EXP-PAR	34.532	33.350	59.026	39.036	38.017	67.735	49.842	47.960	72.499
GAM-LN	42.237	42.095	66.033	43.308	43.680	73.999	53.392	52.332	82.564

Table 9: Results of VaR, HR VaR and TVaR at the 0.90, 0.95 and 0.99 significance levels using the insurance claim data.

Distribution	$VaR_{0.90}$	HR $VaR_{0.90}$	$TVaR_{0.90}$	$VaR_{0.95}$	HR $VaR_{0.95}$	$TVaR_{0.95}$	$VaR_{0.99}$	HR $VaR_{0.99}$	$TVaR_{0.99}$
Empirical	73.962	72.171	97.610	80.134	86.026	118.393	136.281	131.852	144.950
BZTP-EXP	71.412	70.202	95.056	79.395	77.769	117.188	135.514	134.878	144.874
ZTP-EXP	68.050	68.040	92.480	74.640	75.588	115.360	132.159	130.112	138.956
CNB-EXP	66.866	65.843	90.753	69.542	67.419	102.250	128.256	129.185	132.874
NB-EXP	59.154	57.898	83.058	68.136	66.539	98.779	120.725	119.522	124.035
CG-EXP	49.434	48.344	73.016	52.333	51.159	83.047	106.584	102.522	111.911
EXP	49.837	47.544	73.376	54.150	53.314	85.843	107.767	104.563	112.232
EXP-PAR	63.596	59.727	86.601	68.204	66.602	100.951	124.633	122.030	129.603
GAM-LN	72.382	70.787	95.247	79.641	78.674	117.382	135.607	136.949	144.891

6 Conclusions

In this paper, we have developed a new class of distributions by compounding random Poisson sum with exponential random variables. We have discussed different properties of this class of model and derived the statistical inference of the unknown parameters. Furthermore, we also proposed a bivariate version model with three parameters and derive different properties. It is observed that the MLEs of the unknown parameters cannot be obtained in explicit form. We have proposed to use EM type algorithm to compute the MLEs. The proposed EM type algorithm avoids solving three non-linear equations. Further, we have analyzed one real insurance data set and it is observed that the BZTP-EXP model provides a good candidate fit when compared to ZTP-EXP, BCNB-EXP, CNB-EXP, CG-EXP, EXP, EXP-PAR and GAM-LN models. It will be interesting to obtain several properties of this new family of distributions in multivariate case. More work is needed in that direction.

Acknowledgements: We would like to appreciate the constructive comments by an associate editor and anonymous referees which improved the quality and the presentation of our results.

Appendices:

Here, we present some derivations for CNB-EXP model based on combining the exponential and negative binomial distributions as well as descriptions for the Pseudo codes for EM type algorithms for obtaining the MLEs of the unknown parameters under ZTP-EXP and BZTP-EXP models.

Appendix A: CNP-EXP Model

Assume that N is negative binomial (NB) random variable with parameters r and ν. It is easily checked that L is NB distribution with parameters r and $\eta = \nu p$, where $p = P(I_i = 1) = 1 - P(I_i = 0), i = 1, 2, \ldots, N$ with PMF

$$P(L = l) = \binom{l-1}{r-1} \eta^r (1-\eta)^{l-r}, l = r, r+1, \ldots. \tag{37}$$

The joint PDF of CNB-EXP of X and L is given by

$$f_{X,L}(x, l) = \binom{l-1}{r-1} \eta^r (1-\eta)^{l-r} \cdot \frac{\beta e^{-\beta x} (\beta x)^{l-1}}{\Gamma(l)}, \; x > 0, \quad l = r, r+1, \ldots,$$

where $\{X_i, i = 1, 2, \ldots\}$ is a sequence of i.i.d. exponential with parameter β and L follows NB distribution.

The CDF of X is obtained to be

$$\begin{aligned} P(X \le x) &= \sum_{l=r}^{\infty} P(X \le x, L = l) \\ &= \left(\frac{\eta}{1-\eta}\right)^r \sum_{l=r}^{\infty} \binom{l-1}{r-1} (1-\eta)^l \frac{\Gamma_{\beta x}(l)}{\Gamma(l)}. \end{aligned}$$

Therefore the marginal density of X can be expressed as

$$f_X(x) = \left(\frac{\eta}{1-\eta}\right)^r \beta(1-\eta) e^{-\beta x} \sum_{l=r}^{\infty} \binom{l-1}{r-1} \frac{[\beta(1-\eta)x]^{l-1}}{\Gamma(l)}, \qquad x > 0.$$

That is, the PDF of X is an infinite mixture of gamma densities with NB weights. As a result of that, the conditional probability and expectation of L given $X = x$ are readily obtained as

$$P(L = l \mid X = x) = \frac{\binom{l-1}{r-1} \frac{[\beta(1-\eta)x]^l}{\Gamma(l)}}{\sum_{l=r}^{\infty} \binom{l-1}{r-1} \frac{[\beta(1-\eta)x]^{l-1}}{\Gamma(l)}}.$$

The conditional expectation of M given $X = x$ is obtained as

$$E(L \mid X = x) = \frac{\sum_{l=r}^{\infty} l \binom{l-1}{r-1} \frac{[\beta(1-\eta)x]^l}{\Gamma(l)}}{\sum_{l=r}^{\infty} \binom{l-1}{r-1} \frac{[\beta(1-\eta)x]^l}{\Gamma(l)}}.$$

Arguments similar to those in Section 3, we immediately obtain the joint PDF of (X, N) as

$$P(S_L \le x, N \le n) = \sum_{k=1}^{n} \sum_{l=r}^{\infty} \frac{\Gamma_{\beta x}(l)}{\Gamma(l)} \binom{k}{l} p^l (1-p)^{k-l} \binom{k-1}{r-1} \nu^r (1-\nu)^{k-r}.$$

This turns out to be

$$P(S_L \le x, N = n) = \sum_{l=r}^{\infty} \frac{\Gamma_{\beta x}(l)}{\Gamma(l)} \binom{n}{l} p^l (1-p)^{n-l} \binom{n-1}{r-1} \nu^r (1-\nu)^{n-r}.$$

The joint PDF of (S_L, N) can be derived easily.

Appendix B: Pseudo Code of the EM Type Algorithm for ZTP-EXP

- Given IT an integer, denoting the maximum number of iterations allowed. Set $\epsilon = 10^{-5}$, the tolerance limit needed for the stopping the iteration process, that is, if the absolute difference between the two successive log-likelihood values is less than ϵ, then we stop the iteration. Also, set n to denote the sample size.
- Read the data vector $\{x_1, \ldots, x_n\}$.
- Set the initial values $\beta^{(0)}$ and $\theta^{(0)}$ of β and θ, respectively.

  For (j in 1 to IT){

    For (i in 1 to n) {

Find the missing value $l_i^{(j)}$ as follows:

If $\frac{P(L^\star = l+1 \mid x_i, \beta^{(j)}, \theta^{(j)})}{P(L^\star = l \mid x_i, \beta^{(j)}, \theta^{(j)})} < 1$ then $l_i^{(j)} = \min\{l\} \in \{1, 2, \ldots\}$.

}

- Compute $\beta^{(j+1)} = \frac{\sum_{i=1}^n l_i^{(j)}}{\sum_{i=1}^n x_i}$ and $\theta^{(j+1)}$ as a solution of (20), by using bisection method.
- Compute the log-likelihood function $l(\beta^{(j+1)}, \theta^{(j+1)})$.
- If $|l(\beta^{(j+1)}, \theta^{(j+1)}) - l(\beta^{(j)}, \theta^{(j)})| < \epsilon$, we stop the iteration, else we continue until convergence.

}

Appendix C: Pseudo Code of the EM Type Algorithm for BZTP-EXP

- Given IT an integer, denoting the maximum number of iterations allowed. Set $\epsilon = 10^{-5}$, the tolerance limit needed for the stopping the iteration process, that is, if the absolute difference between the two successive log-likelihood values is less than ϵ, then we stop the iteration. Also, set n to denote the sample size.
- Read the data vector $\{(x_1, v_1), \ldots, (x_n, v_n)\}$.
- Set the initial values $\beta^{(1)}$, $\lambda^{(1)}$ and $p^{(1)}$ of β, λ, and p, respectively.

For (j in 1 to IT){

For (i in 1 to n) {

Find the missing value $l_i^{(j)}$ as follows:

If $\frac{P(L = l+1 \mid V = v_i, p^{(j)})}{P(L = l \mid V = v_i, p^{(j)})} < 1$ then $l_i^{(j)} = \min\{l\} \in \{1, 2, \ldots\}$.

}

- Compute $\beta^{(j+1)} = \frac{\sum_{i=1}^n l_i^{(j)}}{\sum_{i=1}^n x_i}$, $\lambda^{(j+1)}$ and $p^{(j+1)}$ as a solution of (35) and (36) respectively, by using the bisection method.
- Compute the log-likelihood function $l(\beta^{(j+1)}, \lambda^{(j+1)}, p^{(j+1)})$.
- If $|l(\beta^{(j+1)}, \lambda^{(j+1)}, p^{(j+1)}) - l(\beta^{(j)}, \lambda^{(j)}, p^{(j+1)})| < \epsilon$, we stop the iteration, else we continue until convergence.

}

References

[1] Andrews, D. F. and A. M. Herzberg. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*, New York: Springer-Verlag.

[2] Asgharzadeh A., Esmaily L., and Nadarajah S.(2013). Approximate MLEs for the location and scale parameters of the skew logistic distribution, *Statistical Papers*, 54, 391-411.

[3] Asgharzadeh A., Esmaily L., and Nadarajah S. and Shih S. H. (2013): A generalized skew logistic, *REVSTAT Statistical Journal*, 11(3), 317-338.

[4] Asgharzadeh, A., Esmaeili, L., and Nadarajah, S. (2016). Balakrishnan skew logistic distribution, *Communication in Statistics-Theory & Methods*, 45(2), 444-464.

[5] Azzalini A. (1985). A Class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, 12(2), 171-178.

[6] Bakouch, H. S., Ristic, M. M., Asgharzadeh, A., Esmaily, L., and Al-Zahrani, B. M. (2012). An exponentiated exponential binomial distribution with application, *Statistics and Probability Letters*, 82, 1067-1081.

[7] Barreto-Souza, W. (2012). The skew-normal distribution and related multivariate families, *Scandinavian Journal of Statistics*, 109, 130-145.

[8] Bernardi, M., Bottone, M. and Petrella, L. (2018). Bayesian quantile regression using the skew exponential power distribution, *Computational Statistics and Data Analysis*, 126, 92-111.

[9] Bhati D. and Ravi S. (2018). On generalized log-Moyal distribution: A new heavy tailed size distribution, *Insurance: Mathematics and Economics*, 79, 247-259.

[10] Birnbaum, Z. and Saunders, S. (1969). A new family of life distributions, *Journal of Applied Probability*, 6(2), 319-327 .

[11] Chakraborty, S., Hazarika, P. J., and Ali, M. M. (2012). A new skew logistic distribution and its properties, *Pakistan Journal of Statistics*, 28(4), 513-524.

[12] Dutang, C., and Charpentier, A. (2019). CASdatasets: Insurance datasets, available at http://dutangc.free.fr/pub/RRepos/web/CASdatasets-index.html.

[13] Grun, B. and T. Miljkovic (2019). Extending composite loss models using a general framework of advanced computational tools, *Scandinavian Actuarial Journal*, 2019(8), 642-660.

[14] Hashemi, F., Naderi, M., and Jamalizadeh, A. (2019). Normal mean-variance Lindley Birnbaum-Saunders distribution, *Statistics and Interface*, 12, 585-597.

[15] Kozubowski, T.J., and Panorska, A.K. (2005). A mixed bivariate distribution with exponential and geometric marginals, *Journal of Statistical Planning and Inference*, 134, 501-520.

[16] Kundu, D. (2014). Geometric skew normal distribution, *Sankhya* B76(2), 167-189.

[17] Kundu, D. and V. Nekoukhou. (2018). Univariate and bivariate geometric discrete generalized exponential distribution, *Journal of Statistical Theory and Practice*, 12(3), 595-614.

[18] Louis, T. A. (1982). Finding the observed information matrix when uisng the EM algorithm, *Journal of the Royal Statistical Society*, Ser. B 44, 226-233.

[19] Ma, Y. and Genton, M.G. (2004). A Flexible class of skew symmetric Distribution, *Scandinavian Journal of Statistics*, 31, 259-468.

[20] Mahmoudi, E. and Jafari, A. A. (2012). Generalized exponential–power series distributions, *Computational Statistics and Data Analysis*, 56, 4047–4066.

[21] McNeil, A. J., Frey, R., and Embrechts, P. (2010). Quantitative risk management: concepts, techniques, and tools, Princeton university press

[22] Miljkovic, T. and Grün, B. (2016). Modeling loss data mixtures of distributions,*Insurance: Mathematics and Economics*, 70, 387-396.

[23] Naderi, M., Hashemi, F., Bekker, A., and Jamalizadeh, A. (2020). Modeling right-skewed financial data streams: A likelihood inference based on the generalized Birnbaum–Saunders mixture model, *Applied Mathematics and Computation*, 376, 125109.

[24] Raqab, M. Z., Kundu, D., and Fahimah A. F. (2019). Compound zero-truncated Poisson normal distribution and its Applications, *Communications in Statistics-Theory and Methods*, To Appear. https://doi.org/10.1080/03610926.2019.1679182.

[25] Revfeim, K. J. A. (1984). An initial model of the relationship between rainfall events and daily rainfalls, *Journal of Hydrology*, 75 (1–4), 357–364.

[26] Rodrigues, C. G., Louzada, F., and Ramos, P. L. (2018). Poisson–exponential distribution: different methods of estimation, *Journal of Applied Statistics*, 45(1), 128-144.

[27] Tomarchio, S. D. and Punzo, A. (2020). Dichotomous unimodal compound models: application to the distribution of insurance losses, *Journal of Applied Statistics*, https://doi.org/10.1080/02664763.2020.1789076.

[28] Wahed A. S. and Ali M. M.(2001). The skew logistic distribution, *Journal of Statistical Research*, 35(2), 71-80.